DATASET BRIEF

Proteomic analysis of germinating urediniospores of *Phakopsora pachyrhizi*, causal agent of Asian soybean rust

Douglas G. Luster[1], Michael B. McMahon[1], Melissa L. Carter[1], Laurie L. Fortis[2] and Alberto Nuñez[2]

[1] Foreign Disease-Weed Science Research Unit, US Department of Agriculture, Agricultural Research Service, MD, USA
[2] Eastern Regional Research Center, US Department of Agriculture, Agricultural Research Service, PA, USA

Phakopsora pachyrhizi is an obligate pathogen that causes Asian soybean rust. Asian soybean rust has an unusually broad host range and infects by direct penetration through the leaf cuticle. In order to understand the early events in the infection process, it is important to identify and characterize proteins in *P. pachyrhizi*. Germination of the urediniospore is the first stage in the infection process and represents a critical life stage applicable to studies with this obligate pathogen. We have applied a 2-DE and MS approach to identify 117 proteins from the National Center of Biotechnology Information nonredundant protein database and a custom database of Basidiomycota EST sequences. Proteins with roles in primary metabolism, energy transduction, stress, cellular regulation and signaling were identified in this study. This data set is accessible at http://world-2dpage.expasy.org/repository/database = 0018.

Received: July 1, 2009
Revised: May 4, 2010
Accepted: July 6, 2010

Keywords:
2-DE / Fungal Proteins / MALDI-TOF/TOF / Microbiology / *Phakopsora Pachyrhizi* / Urediniospores

Phakopsora pachyrhizi is the causal agent of Asian soybean rust (ASR), a significant foliar disease of soybeans (*Glycine max*) and other legume crops that occur in most soybean growing regions throughout the world, except Europe. The pathogen has spread rapidly from its origin in Asia to Southern Africa and South America over the past 5 years, entering the US in 2004 and raising concern to US soybean producers [1, 2].

P. pachyrhizi is an obligate pathogen, spread primarily by the airborne dispersal of urediniospores [3, 4]. The rate of disease progression is dependent on temperature and available moisture [5]. *P. pachyrhizi* has an unusually wide host range for a rust fungal pathogen [6, 7], and is able to infect a large variety of plants in the Leguminoseae family. *P. pachyrhizi* is unique in that it enters the plant by direct penetration through the cuticle using enzymatic digestion and physical force [6, 8]. Most rusts enter through stomata using physical force, whereas other fungi enter by enzymatic digestion of the cell wall [9–12]. Urediniospores are the only life stage amenable to studies of the genome, transcriptome and proteome, representing the only developmental stage that can be generated in sufficient quantities free of host material. Although 48 567 *P. pachyrhizi* ESTs have been sequenced and deposited into GenBank, little is currently known about the *P. pachyrhizi*–soybean interaction at the molecular level.

In order to understand the early events in the infection process, it is important to identify and characterize proteins in *P. pachyrhizi* urediniospores. In this study, we have focused on urediniospore germination, which represents a critical stage in the infection process for

Correspondence: Dr. Douglas G. Luster, Foreign Disease-Weed Science Research Unit, USDA, ARS, 1301 Ditto Avenue, Ft. Detrick, MD 21702, USA
E-mail: Doug.Luster@ars.usda.gov
Fax: +1-301-619-2880

Abbreviations: NCBI, National Center of Biotechnology Information; **NCBInr,** nonredundant NCBI; **PFF,** peptide fragmentation fingerprint

identification of genes involved in early infection and interaction with host legumes. To address this challenge, we have applied 2-DE and MS to identify predominantly soluble proteins present during the germination phase of ASR. Our analysis has identified a variety of fungal proteins with demonstrated roles in cell biosynthesis/energy/metabolism, regulatory/signaling, stress responses and infection processes.

Urediniospores of *P. pachyrhizi* isolate Taiwan 72-1 were used in four independent replicate germination and protein extraction experiments. Approximately, 300 mg of urediniospores were germinated under sterile conditions for 18 h in a 23×28 cm Pyrex dish containing 300 mL of water, supplemented with 50 μg/mL of ampicillin and 25 μg/mL of streptomycin. Germinated urediniospores were collected by filtration through a Buchner funnel with Whatman #1 filter paper. All extractions were performed at the USDA biosafety level 3 Plant Pathogen Containment Greenhouse Facility at Ft. Detrick, MD, under the appropriate USDA-Animal Plant and Health Inspection Service permit.

Samples were ground to a fine powder in liquid nitrogen and suspended on ice in five volumes of extraction buffer (25 mM Tris, pH 7.0, 5 mM EDTA, 3 mM DTT, 5% glycerol v/v and 10 μL/mL of protease inhibitor cocktail (Sigma, St. Louis, MO)). The suspension was mixed thoroughly and centrifuged for 5 min at $500 \times g$ to remove intact spores and spore wall fragments, followed by centrifugation at $10\,000 \times g$ for 15 min at 4°C. The extracts were then filtered through a 0.45 μm syringe filter unit in a sterile hood into a sterile vial and surface decontaminated with 10% bleach solution for removal from the biosafety level 3 facility. Samples were dialyzed against a buffer of 10 mM Tris, pH 7.0, 1 mM EDTA and 3 mM DTT for 6 h. Protein quantification was performed using the Markwell assay [13], with BSA as a standard.

Protein samples were precipitated using the 2-D clean up kit (GE Healthcare, Piscataway, NJ) and 200–400 μg were resuspended in DeStreak Rehydration buffer (GE Healthcare) containing 1% ampholytes v/v (Invitrogen, Carlsbad, CA). Solublized proteins were applied to a 13 cm 3–10 nonlinear IPG Dry Gel Strip (GE Healthcare). IEF was carried out on an Ettan IPGphor system (GE Healthcare) at 20°C at 50 μAmp/strip for a total of 20 000 Vh. For second dimension, strips were equilibrated in 50 mM Tris, pH 7.0, 6 M Urea, 30% v/v glycerol, 2% SDS and 10 mg/mL DTT for 10 min, followed by 10 min in equilibration buffer containing 25 mg/mL iodoacetamide and run on 4–12% BisTris gels. Gels were stained with Simply Blue Safe Stain (Invitrogen) and images documented on an Alpha Innotech Fluor-Chem 8900. Experiments are described in the MIAPE document at http://miapegeldb.expasy.org/experiment/51/. Figure 1 shows a 2-D gel from one representative replicate. Numbered protein spots depict samples that were identified in that experiment. Approximately, 120 spots were excised from each gel, for a total of 480 spots picked from four gels, each loaded with an extract from a replicated germination

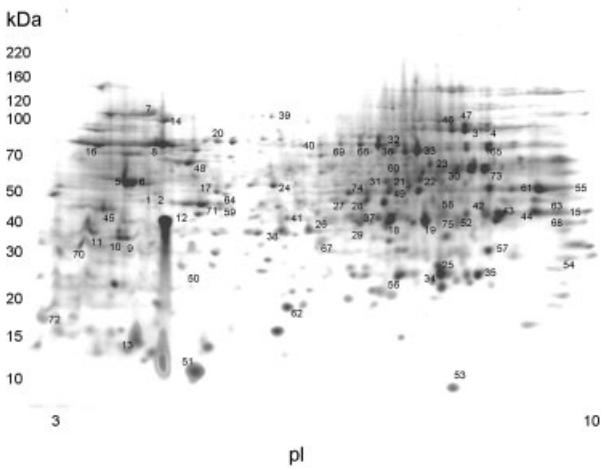

Figure 1. 2-DE of *P. pachyrhizi* urediniospores separated on 13 cm, pH 3–10 nonlinear IPG strip (IEF) followed by 4–12% SDS-PAGE. Numbered protein spots depict samples that were identified in this experiment.

experiment. Selected protein spots were destained, trypsin-digested and peptides extracted using C18 filter tips. Peptides were spotted onto a MALDI plate in triplicate resulting in 1440 samples analyzed.

All spectra were obtained using the Applied Biosystems 4700 Proteomics Analyzer mass spectrometer (Framingham, MA) in the positive reflectron mode with a 200 Hz Nd-YAG 355 nm laser. Trypsin-digested proteins were subjected to MALDI-TOF/TOF analysis. Spectra acquired between the range of 800 and 4000 Da in MS mode were obtained through the averaging of 1000 spectra, and 2000 spectra in the MS/MS mode. Up to seven ions were selected using a signal-to-noise cutoff of 15 for low-intensity peptide signals, and 30 for high-intensity peptides, excluding all common trypsin autolysis peaks and common keratin contaminants. These ions were subjected to PSD or CID with air as the collision gas at approximately 1×10^6 Torr, and a 1 keV acceleration voltage. Instrument default calibration to a mass tolerance of ± 50 ppm was achieved using a minimum signal-to-noise ratio of 10 for MS calibration with the 4700 Standard Peptide Calibration Mixture (Applied Biosystems), used for the conversion of time-of-flight to mass (Da) for the monoisotopic ions, $[M+H]^+$. The MS/MS TOF calibration was achieved to a m/z ± 0.3 tolerance from the PSD of Glu^1-fibrinopeptide B fragments using a minimum signal-to-noise ratio of 10.

All spectra acquired were processed using default calibration or internal standard calibration from two to three trypsin autolysis peaks typically in each sample after digestion including peaks at 842.51, 2211.10 and 3246.11 Da. Spectra were subsequently analyzed using the MASCOT [14] search engine-associated GPS Explorer program (Applied Biosystems) against a public database and a custom sequence database of EST sequences in FASTA format. Peptide mass fingerprints and MS/MS of selected peptides,

commonly known as the peptide fragmentation fingerprints (PFF), were combined for database analyses. The analyses included the variable modifications methionine oxidation, *N*-terminal glutamine forming an internal cyclic lactam to become pyrrolidone carboxylic acid and carbamidomethylation of cysteine residues from the reduction and alkylation of proteins with a mass tolerance of 0.5 ppm. All proteins reported from database searches of putative peptide sequences are within a 95% confidence interval. The full data set from these analyses is included in the Supporting Information.

Databases included the National Center of Biotechnology Information (NCBI) nonredundant (NCBInr) protein entries and a subset of the NCBI EST database compiled in March 2010. The EST subset was obtained by searching NCBI using the keyword ''Basidiomycota'' representing 615 150 fungal sequences from genera and species with established EST projects.

The EST database contains 48 567 *P. pachyrhizi* entries with high degree of redundancy. Two decoy databases were also created by the randomization of each sequence in these databases to maintain the composition and distribution of sequence lengths imitating true entries, and searched to provide a false-positive rate for each database. All databases were searched using MASCOT version 2.2.06 with GPS Analysis software version 3.6 (Applied Biosystems).

Ion scores for a PFF are based on a probability-based molecular weight search score [15], $-10^{*}\mathrm{Log}\,(P)$, where the probability (P) that the match observed is a random event. Protein scores are based on the sum of the ion scores for the selected PFF of the sample. The protein score thresholds for the NCBI and EST databases respectively, were 83 and 78, where a match score greater than these thresholds is considered significant within a 0.05 probability. Ion scores with an *E*-value of 0.02 with the database are considered to indicate extensive similarity or identity with the database entry. Databases were validated by reanalyzing all samples using a decoy database composed of randomized entries of the corresponding original database. Of the 1495 samples analyzed, the randomized NCBI showed false-positive matching scores for 15 samples, and the randomized EST database had no positive matches.

Combined MS and MSMS analyses resulted in 393 putative identities within the established threshold limits (Supporting Information Table S1). Of these, 127 are identified in the NCBInr database and 266 in the EST database. Putative peptide sequences for each protein are listed in Supporting Information Table S1. For this study, we identified 117 different protein spots with the highest scoring sequence matches and greatest peptide coverage. Putative proteins identified in the EST database are summarized in Tables 1 and 2. Of the 117 ESTs, 101 (86%) are *P. pachyrhizi*, 13 (11%) related rust species, 2 (2%) *Postia placenta* and 1 (1%) *Laccaria bicolor*. Table 2 summarizes a subset of 33 entries identified by a single peptide. GenBank EST submissions are not annotated and contain short single-pass sequences; therefore, the possibility of identifying an EST with a single peptide is increased on any one analysis. Coupled with the fact that no false positives occurred in the EST decoy database, and 28 (85%) of the identities are to *P. pachyrhizi*, we feel that these data are biologically relevant to the study. A total of 33 different protein spots are identified in the NCBI database search (Supporting Information Table S2). All NCBI proteins were confirmed in the EST data set as well; those identified represent a group of well-characterized, highly conserved proteins routinely found in other proteomic studies of fungal (31) and plant (2) species.

Putative protein identities were assigned to ESTs in Tables 1 and 2 based on the BLAST searches against the NCBInr protein database, using ORFs matching the peptides for each accession. Proteins were categorized using Uni-Prot Protein Knowledge Database [16], and placed into eight functional categories using the Munich Information Center for Protein Sequences (http://mips.gsf.de/) classification system. The eight functional categories were determined to be a percentage of the total number of proteins for each grouping; the proteins with multiple functions were given secondary classification. Figure 2 shows the primary and secondary functional categories of 117 proteins in Tables 1 and 2. As expected, the category encompassing the highest percentage of isolated proteins was cell biosynthesis/energy/metabolism with 40 proteins (34%). Numerous proteins involved in energy production were linked to glycolysis and the citric acid cycle: aldehyde reductase, enolase, phosphoglycerate kinase, phosphoenolpyruvate kinase, glyceraldehyde-3-phosphate dehydrogenase, triose phosphate isomerase, L-malate dehydrogenase, citrate synthase and phosphoglycerate mutase. Primary metabolism proteins for the pentose phosphate pathway were also prevalent: transaldolase, phosphoenolpyruvate carboxykinase, aspartate aminotransferase, phosphoglucomutase and transketolase. Additional proteins with primary and secondary cell biosynthetic functions were identified and could be attributed to glycosylation, ATP transport, amino acid synthesis, electron transport, nitrogen assimilation and fatty acid metabolism.

Stress–response proteins accounted for 19 proteins (16%) within the functional categories. Nine were directly related to the large family of chaperone and HSPs. HSPs play an important role in monitoring and recycling the cell's proteins during rapid cell division and breaking cell dormancy [17]. Three well-characterized enzymes (catalase, peroxidase and manganese superoxide dismutase) are represented. These enzymes are responsible for removal of the oxygen-free radicals and hydrogen peroxide that rapidly accumulate during spore germination. Twenty-one proteins (18%) were directly related to translation and protein synthesis, including arginyl-tRNA synthetase, initiation factor 4A and several elongation factors. A group of related proteins involved in protein degradation and turnover were also identified: ubiquitin-activating enzyme E1 and ubiquitin-conjugating enzyme E2, proteasome regulatory

Table 1. Proteins identified using the custom EST database

Spot	Putative protein	Acc.	Pep.	MASCOT score	*E*-value	p*I*	MW	Function	Local
3	*P. pachyrhizi* aconitate hydratase	gil120512044	3	359	4.40E−30	8.87	26 350.6	cb/en/met	Cyto
4	*P. pachyrhizi* aconitate hydratase	gil120512044	3	323	1.80E−26	8.87	26 350.6	cb/en/met	Cyto
23	*P. pachyrhizi* phosphoglucomutase	gil120523315	2	172	2.20E−11	8.66	31 726.6	cb/en/met	Cyto
32	*P. pachyrhizi* transketolase	gil120519330	4	577	7.00E−52	9.12	30 034.4	cb/en/met	Cyto
33	*P. pachyrhizi* transketolase	gil120519330	4	498	5.50E−44	9.12	30 034.4	cb/en/met	Cyto
63	*P. pachyrhizi* aspartate amino transferase[a)]	gil120522279	3	214	1.40E−15	8.12	29 689.8	cb/en/met	Cyto
22	*P. pachyrhizi* phosphogluconate dehydrogenase[a)]	gil120523598	3	328	5.50E−27	9.41	28 595.8	cb/en/met	Cyto
26	*P. pachyrhizi* inorganic pyrophosphatase[a)]	gil120523830	6	385	1.10E−32	6.41	30 848.6	cb/en/met	Cyto
30	*P. pachyrhizi* methylcitrate dehydratase	gil120503921	3	134	1.40E−07	8.88	21 555.3	cb/en/met	Cyto
57	*P. pachyrhizi* triose phosphate isomerase[b)]	gil120499172	6	418	5.50E−36	8.78	25 886.6	cb/en/met	Cyto
58	*P. pachyrhizi* phosphoglycerate kinase[a)]	gil120512620	4	266	8.80E−21	9.18	29 443.8	cb/en/met	Cyto
66	*P. pachyrhizi* phosphoenolpyruvate kinase	gil120534934	2	120	3.50E−06	9.69	29 593.2	cb/en/met	Cyto
75	*P. pachyrhizi* L-malate dehydrogenase[a)]	gil120523458	2	101	2.80E−04	8.54	29 381.8	cb/en/met	Mito
18	*P. pachyrhizi* transaldolase[a)]	gil120510700	3	302	2.20E−24	8.96	30 933.1	cb/en/met	Cyto
19	*P. pachyrhizi* transaldolase[a)]	gil120510700	4	422	2.20E−36	8.96	30 933.1	cb/en/met	Cyto
29	*P. pachyrhizi* glutamine synthetase	gil120508497	3	396	8.80E−34	5.71	28 543.2	cb/en/met	Cyto
42	*U. appendiculatus* glyceraldehyde-3-phosphate dehydrogenase[a)]	gil120428168	3	330	3.5E−27	8.96	28 667.8	cb/en/met	Cyto
43	*U. appendiculatus* glyceraldehyde-3-phosphate dehydrogenase[a)]	gil169739158	3	181	2.80E−12	7.82	22 682.7	cb/en/met	Cyto
44	*Pu. triticina* glyceraldehyde-3-phosphate dehydrogenase[a)]	gil282811836	4	441	2.80E−38	5.97	29 176.5	cb/en/met	Cyto
52	*P. pachyrhizi* aldehyde reductase[a)]	gil120522626	3	296	8.80E−24	6.60	28 447.1	cb/en/met	Cyto
64	*P. pachyrhizi* prolidase	gil83922287	3	109	4.40E−05	5.63	25 833.1	cb/en/met	Cyto
65	*Melampsora larici-populina* fumarate reductase[a)]	gil254151033	4	109	1.80E−09	6.66	28 090.1	cb/en/met Trpt	Cyto
76	*P. pachyrhizi* phosphoglycerate mutase[a)]	gil120526000	2	243	1.80E−18	9.20	25 796.6	cb/en/met	Cyto
77	*P. pachyrhizi* short-chain dehydrogenase	gil120507723	2	306	8.80E−25	9.02	31 145.5	cb/en/met	Cyto
78	*P. triticina* NADH dehydrogenase (ubiquinone)[a)]	gil282827075	2	205	1.10E−14	8.13	27 635.8	cb/en/met	Mito
83	*P. pachyrhizi* methionine adenosyl-transferase[a)]	gil120518043	2	85	1.00E−02	6.35	30 305.7	cb/en/met	Cyto
84	*Po. placenta* valosin-containing protein	gil197139912	3	295	1.10E−23	6.84	24 052.4	cb/en/met	Cyto
108	*P. pachyrhizi* NAPD-dependent mannitol dehydrogenase	gil120495319	3	306	8.80E−25	9.02	26 846.5	cb/en/met Stress	Cyto
117	*P. pachyrhizi* fasciclin domain family protein	gil120515828	2	109	4.40E−05	6.28	23 848.4	cb/en/met	Extc
1	*P. pachyrhizi* initiation factor 4a[b)]	gil120503558	2	146	7.00E−09	4.95	27 518.4	Prosyntur	Cyto
2	*P. pachyrhizi* initiation factor 4a[b)]	gil120524582	4	428	5.50E−37	8.89	29 254.4	Prosyntur	Cyto
15	*P. pachyrhizi* elongation factor 1-α[a)]	gil120520943	5	436	8.80E−38	8.60	28 951.9	Prosyntur	Cyto
17	*P. pachyrhizi* RNA helicase[a)]	gil120499251	5	353	1.80E−29	5.50	25 412.2	Prosyntur Trscp	Cyto
28	*P. pachyrhizi* elongation factor γ	gil120506356	2	174	1.40E−11	7.21	31 130.9	Prosyntur	Nuc
39	*P. pachyrhizi* ubiquitin-activating enzyme E1	gil120518512	2	80	9.40E−03	5.10	29 663.9	Prosyntur cb/en/met	Nuc

Table 1. Continued

Spot	Putative protein	Acc.	Pep.	MASCOT score	*E*-value	p*I*	MW	Function	Local
46	*P. pachyrhizi* translation elongation factor 2[a)]	gil120526577	2	112	2.80E−77	7.66	27 010.2	Prosyntur	Cyto
47	*P. placenta* translation elongation factor 2[a)]	gil220023712	3	210	3.50E−15	9.13	76 496.1	Prosyntur	Cyto
59	*P. pachyrhizi* protein disulfide isomerase	gil120509243	7	312	2.20E−25	4.90	30 417.2	Prosyntur cb/en/met	ER
69	*P. pachyrhizi* arginyl-tRNA synthetase[c)]	gil120526328	2	108	5.80E−05	5.57	23 489.9	Prosyntur	Mito
70	*P. pachyrhizi* elongation factor β[c)]	gil120535387	3	262	2.20E−20	4.64	25 621.7	Prosyntur	Cyto
87	*P. pachyrhizi* aspartic peptidase	gil120505360	2	244	1.40E−18	5.56	30 707.7	Prosyntur	ER
88	*P. pachyrhizi* proteasome regulatory particle	gil120523757	2	105	1.10E−04	6.09	23 687.2	Prosyntur	Cyto
114	*P. pachyrhizi* 40S ribosomal protein	gil120422127	2	133	1.80E−07	9.28	25 055.0	Prosyntur	Cyto
115	*P. pachyrhizi* 60S ribosomal protein L12	gil120496396	2	179	4.40E−12	8.48	24 618.1	Prosyntur	Cyto
9	*P. pachyrhizi* 14-3-3[c)]	gil120520171	3	397	7.00E−34	6.78	30 747.8	reg/sig Trscp	Cyto
10	*P. pachyrhizi* 14-3-3[c)]	gil120523695	8	673	1.80E−61	5.09	30 190.2	reg/sig Trscp	Cyto
11	*P. pachyrhizi* 14-3-3[c)]	gil120523695	5	384	2.80E−33	5.09	30 190.2	reg/sig Trscp	Cyto
50	*P. pachyrhizi* ρ GDP dissociation inhibitor[c)]	gil120517577	6	351	2.80E−29	5.52	32 467.3	reg/sig Struct	Cyto
67	*P. pachyrhizi* proteasome regulatory particle[a)]	gil120523757	2	114	1.40E−05	6.09	23 687.2	reg/sig	Cyto
94	*P. pachyrhizi* voltage-dependent ion selective channel	gil120493672	3	142	2.20E−08	9.42	29 212.3	reg/sig	PM
8	*P. pachyrhizi* HSP60[a)]	gil120499849	3	197	7.00E−14	9.9	26 462.2	Stress	Mito
14	*P. pachyrhizi* HSP[a)]	gil120498065	5	434	1.40E−37	7.97	25 660.7	Stress	Mito
20	*P. pachyrhizi* HSP70[b)]	gil120504670	5	664	1.40E−60	5.68	24 819.1	Stress	Cyto
31	*P. pachyrhizi* α ketoglutarate dioxygenase	gil120506411	2	80	3.90E−02	7.16	17 819.2	Stress	Cyto
		gil83922374		90	3.20E−03	8.65	21 621.2	Stress	Cyto
35	*P. pachyrhizi* manganese superoxide dismutase	gil120515897	4	279	4.40E−22	9.34	29 396.2	Stress cb/en/met	Mito
37	*P. pachyrhizi* small HSP	gil83922383	2	141	2.80E−08	8.12	23 988.3	Stress	Cyto
48	*U. appendiculatus* 60 kDa chaperonin[c)]	gil169736001	4	358	5.50E−30	5.48	25 674.7	Stress	Cyto
62	*P. pachyrhizi* small HSP	gil120513061	3	401	2.80E−31	8.83	24 718.4	Stress	Cyto
73	*P. pachyrhizi* catalase	gil120527236	5	551	2.80E−49	6.74	57 053.8	Stress reg/sig	Pero
79	*P. pachyrhizi* HSP88	gil120498065	2	128	5.50E−07	7.97	25 660.7	Stress	Cyto
90	*P. pachyrhizi* peroxiredoxin	gil120493268	2	206	8.80E−15	8.99	22 575.0	Stress	Cyto
91	*M. larici-populina* HSP 81-2	gil254164737	2	157	7.00E−10	5.49	29 228.1	Stress	Cyto
93	*P. pachyrhizi* HSP sks2	gil120528518	2	274	5.50E−22	5.02	17 616.9	Stress	Cyto
118	*P. pachyrhizi* glutathione-*S*-transferase	gil120493978	2	235	1.10E−17	8.71	28 560.8	Stress	Cyto
41	*P. pachyrhizi* mannose phosphate guanyltransferase[a)]	gil120528414	3	287	7.00E−23	6.39	23 076.3	Struct cb/en/met	Cyto
71	*P. pachyrhizi* actin[a)]	gil120536058	3	482	2.20E−42	4.96	27 180.6	Struct	Nuc
95	*U. appendiculatus* β tubulin[a)]	gil120425987	2	190	3.50E−13	4.94	28 710.2	Struct	Cyto

Table 1. Continued

Spot	Putative protein	Acc.	Pep.	MASCOT score	*E*-value	p*I*	MW	Function	Local
45	*P. pachyrhizi* endopeptidase	gil120502813	3	206	8.80E−15	6.38	25747.0	Trscp Prosyntur	Cyto
96	*P. pachyrhizi* mismatch recognition protein	gil120536629	3	276	8.80E−22	8.03	12114.1	Trscp	Nuc
5	*P. pachyrhizi* ATP synthase β[b)]	gil120503186	3	346	2.80E−39	4.92	23379.1	Trpt cb/en/met	Mito
6	*P. pachyrhizi* ATP synthase β[b)]	gil120503186	5	531	2.80E−47	4.92	23379.1	Trpt cb/en/met	Mito
7	*P. pachyrhizi* ATP synthase β[b)]	gil120503186	2	439	4.40E−38	4.92	23379.1	Trpt cb/en/met	Mito
24	*P. pachyrhizi* rab GDP dissociation inhibitor[c)]	gil120522716	4	231	2.80E−17	5.64	31164.8	Trpt reg/sig	Cyto
25	*P. pachyrhizi* ran GTP-binding protein[a)]	gil120536717	3	263	1.80E−20	7.36	22308.3	Trpt reg/sig	Nuc
55	*Phaseolus vulgaris* ATP synthase α[a)]	gil171590605	5	219	4.40E−16	9.28	32438.9	Trpt cb/en/met	Mito
68	*P. pachyrhizi* voltage-gated potassium channel subunit	gil120507959	4	155	1.10E−09	9.60	29977.1	Trpt reg/sig	PM
97	*P. pachyrhizi* ADP ribosylation factor	gil83922262	2	153	1.80E−09	8.60	25507.0	Trpt	Golgi
98	*P. pachyrhizi* thioltransferase	gil120526362	3	295	5.50E−19	9.40	25893.6	Trpt	Mito
101	*P. pachyrhizi* nuclear transport factor 2	gil120527820	4	185	1.10E−12	9.49	22152.4	Trpt	Nuc
107	*P. pachyrhizi* ATP synthase delta	gil120505677	2	198	5.50E−14	7.90	23484.1	Trpt cb/en/met	Mito
12	*P. pachyrhizi* unknown protein	gil120528161	4	459	8.80E−27	5.63	18015.7	Unk	Unk
13	*P. pachyrhizi* unknown protein	gil120518981	3	326	4.40E−40	9.01	21897.2	Unk	Unk
49	*P. pachyrhizi* unknown protein	gil120507616	3	283	1.80E−22	7.41	29448.5	Unk	Unk
110	*P. pachyrhizi* unknown protein	gil120516698	2	141	2.80E−08	9.22	31283.9	Unk	Unk

Acc, accession; Pep, peptide count. Local, localization; cb/en/met, cell biosynthesis/energy/metabolism; cyto, cytoplasm; mito, mitochondria; trpt, transport; stress, stress response; extc, extracellular; prosyntur, protein synthesis and turnover; trscrp, transcription; nuc, nucleus; ER, endoplasmic reticulum; reg/sig, regulatory/signaling; struct, structural; stress, stress response; pero, peroxisome; PM, plasma membrane; golgi, golgi apparatus; unk, unknown.

a) Proteins common to Cooper *et al.* [22]

b) Proteins common to both Cooper *et al.* [22] and Rampitsch *et al.* [23]

c) Proteins common to Rampitsch *et al.* [23]

Table 2. Single-peptide identifications from the custom EST database

Spot	Putative protein	Acc.	Peptide	MASCOT Score	*E*-value	p*I*	MW	Function	Local
51	*P. pachyrhizi* conidiation specific protein	gil120499550	LENEEFSSANDR	167	7.00E−11	9.75	21357.3	cb/en/met reg/sig	Cyto
21	*P. pachyrhizi* phosphogluconate dehydrogenase[a)]	gil120497320	DYFGAHTFR	84	5.50E−27	9.41	28595.8	cb/en/met	Cyto
36	*Ph. vulgaris* (phospho)pyruvate hydratase (enolase)	gil171630857	NAGIQIVGDDLTVTNPVR	121	2.80E−06	6.77	24876.9	cb/en/met	Cyto
74	*P. pachyrhizi* acyl Co-A oxidase	gil120498714	VAQEASGSAIEWCGGVGFTR+ Carbamidomethyl C13	158	5.50E−10	7.88	18723.5	cb/en/met	Cyto
38	*P. pachyrhizi* unknown short-chain dehydrogenase	gil120520592	DFGSVDVLICSAGVADNI AAEDYPADR + Carbamidomethyl C10	175	1.10E−11	8.49	27849.1	cb/en/met	Cyto
60	*P. pachyrhizi* phosphoenolpyruvate carboxykinase[a)]	gil120522248	NAPVAQLYEEAIR	113	1.80E−05	10.47	31590.4	cb/en/met	Cyto
61	*P. pachyrhizi* citrate synthase	gil120500208	SIHEELSEIIPER	175	1.10E−11	6.59	31013.7	cb/en/met	Mito
80	*P. pachyrhizi* cytochrome *C* oxidase	gil120504071	MEGHNPFDR	69	4.50E−01	6.82	26476.3	cb/en/met	Mito
81	*P. pachyrhizi* phosphomannomutase	gil120510176	GTFIEFR	91	3.00E−03	9.06	31427.4	cb/en/met	Cyto
82	*P. pachyrhizi* nitrophenylphosphatase	gil120511236	ADLEAKPTYVIESLGDLER	103	1.80E−04	7.70	20274.6	cb/en/met	Cyto
113	*P. pachyrhizi* nucleosome assembly protein	gil120491203	LELDYQIGEDLKDR	99	2.80E−04	4.60	24630.0	cb/en/met	Nuc
116	*P. pachyrhizi* coenzyme A transferase	gil120514294	CVSQVITELAVFDIDR + Carbamidomethyl C1	131	2.80E−07	5.31	27944.9	cb/en/met	Cyto
27	*P. pachyrhizi* elongation factor γ	gil120504672	NIDLSEYSFWR	111	2.80E−05	8.61	25327.6	Prosyntur	Nuc
53	*Laccaria bicolor* ubiquitin	gil170098593	ESTLHLVLR	106	8.80E−05	5.70	8152.4	Prosyntur	Cyto
85	*P. pachyrhizi* ubiquitin conjugating enzyme E2[a)]	gil120504869	YFNVELQGPEGSAFQGGLFK	118	5.50E−06	8.99	22908.0	Prosyntur	Cyto
86	*P. pachyrhizi* serine protease	gil120507764	GVDVYVIDTGINIEHEELEGR	173	1.80E−11	5.46	27849.9	Prosyntur	Cyto
89	*P. pachyrhizi* 20S proteasome subunit[b)]	gil120522775	DNLSAGIIVAGWDSQTK	186	8.80E−13	8.32	27348.4	Prosyntur	Cyto
106	*U. appendiculatus* peptidyl-prolyl cis-trans isomerase	gil120422706	IHYVGTLETK	85	1.10E−02	8.96	21845.2	Prosyntur	Cyto
72	*P. pachyrhizi* calmodulin	gil120504950	EADVDGDGAINYEEFVR	158	5.50E−10	4.17	15864.5	reg/sig cb/en/met	PM
16	*P. pachyrhizi* HSP90[a)]	gil120522584	DMSVLLFETALLTSGF TLDAPQNFAER	81	2.70E−02	4.82	21369.6	Stress	Cyto
34	*P. pachyrhizi* manganese superoxide dismutase	gil120512041	FSQAVSQEFGGLEQLK	170	3.50E−11	9.92	24665.9	Stress cb/en/met	Mito
40	*P. pachyrhizi* fungal peroxidase	gil120508961	DGSFIVFR	87	3.50E−03	5.53	31424.9	Stress cb/en/met	Extc
92	*P. pachyrhizi* Cro RII – HSP70 precursor protein[a)]	gil120523195	AQTWLDENTATATSEDFDEQR	127	7.00E−07	6.56	30514.5	Stress	Cyto
104	*P. pachyrhizi* small HSP	gil83922383	MDVVETEDSFLVTTELPGAK	146	8.80E−09	8.12	23988.3	Stress	Cyto
54	*P. pachyrhizi* 60S ribosomal protein[a)]	gil120524540	GNIGFVFTSGDLK	126	8.80E−07	4.41	27236.0	Struct	Cyto
99	*Ph vulgaris* het-C2	gil171585773	DATQGLLWLTR	93	1.80E−03	9.31	29808.7	Trpt	Cyto
100	*P. pachyrhizi* sec-14 – phosphatidyl inositol transport	gil120518708	LGHLSVEQVHVLK	113	1.80E−05	9.01	32454.8	Trpt	Cyto
102	*P. pachyrhizi* vacuolar ATPase[a)]	gil120499313	AIVQVFEGTSGIDVR	81	2.20E−06	4.87	12065.0	Trpt	PM
56	*P. pachyrhizi* unknown protein	gil120528463	NIVDFNFGTLIR	98	5.70E−04	7.85	18860.6	Unk	Unk
103	*P. pachyrhizi* unknown protein	gil120507720	NTYDDAAQNAYDLQSDNR	153	1.80E−09	6.44	30460.2	Unk	Unk
105	*M. larici-populina* Atg gene	gil254159346	EFDGFLFGFPTR	105	1.40E−04	6.37	27390.8	Unk	Unk

Table 2. Continued

Spot	Putative protein	Acc.	Peptide	MASCOT Score	*E*-value	p*I*	MW	Function	Local
109	*P. pachyrhizi* unknown protein	gi|120498234	SMFLTLNQAGNSGIYNVNR	93	1.70E−03	9.17	23476.6	Unk	Unk
111	*P. pachyrhizi* unknown protein	gi|120527530	QELETVFGSSNDVEVITTVLEK	202	2.20E−14	8.92	26883.9	Unk	Unk
112	*P. pachyrhizi* unknown protein	gi|120500785	VIINVIPQDQHAR	82	2.40E−02	9.10	13635.5	Unk	Unk

Acc, accession; Local, localization; cb/en/met, cell biosynthesis/energy/metabolism; cyto, cytoplasm; reg/sig, regulatory/signaling; mito, mitochondria; nuc, nucleus; prosyntur, protein synthesis and turnover; PM, plasma membrane; stress, stress response; extc, extracellular; struct, structural; trpt, transport; unk, unknown.

a) Proteins common to Cooper *et al.* [22]

b) Proteins common to both Cooper *et al.* [22] and Rampitsch *et al.* [23]

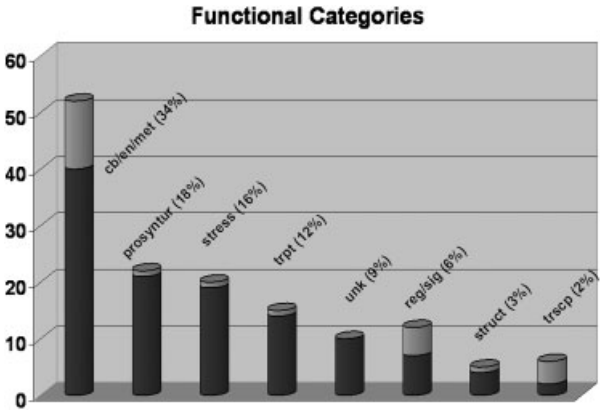

Figure 2. Distribution of identified proteins by functional categories.

particle and 20S proteasome subunit. The E1 and E2 enzymes are responsible for the first steps in the ubiquitination of proteins targeted for degradation in the proteasome [18]. Ten proteins (9%) were defined as unknowns, of which nine matched *P. pachyrhizi* EST ORFs. Two of these unknowns were the most highly abundant protein spots picked from 2-D gels (spots 12 and 13). Highly abundant proteins from other categories included: regulatory/signaling – 14-3-3 [19] and calmodulin [20]; transport – ATP synthase β and GTP-binding protein and structural – actin.

As summarized in Tables 1 and 2, the subcellular localization has been noted for proteins having a putative function. Subcellular localization was assigned based on the database searches. Given the nature of sample preparation and enrichment for soluble fraction, the majority of the proteins (64%) were localized to the cytoplasm. Proteins were also localized to the mitochondria, nucleus, endoplasmic reticulum, membranes, peroxisome and Golgi apparatus.

Few studies have applied a proteomic approach to identify and characterize proteins in an obligate fungal pathogen [21–23]. To our knowledge, this is the first to use a 2-DE approach to identify proteins from germinating rust fungal urediniospores. Our results corroborate a recent study by Cooper *et al.* [22], applying a shotgun approach to identify proteins in ungerminated bean rust (*Uromyces appendiculatus*) urediniospores, with 35 identified proteins in common between the two studies (denoted by superscript [a)] in Tables 1 and 2). Our data set also includes eleven specific proteins in common with those found by Rampitsch *et al.* [23], (denoted by superscript [c)] in Tables 1 and 2), in a 2-DE study of wheat leaves inoculated with *Puccinia triticina* (wheat leaf rust). Five specific proteins were found to be common to all three studies (denoted by superscript [b)] in Tables 1 and 2).

Our study provides a baseline protein inventory for *P. pachyrhizi*, representing the subset of fully soluble proteins present at the time of maximum germling expansion 18 h post-germination. The data set has been deposited into the World-2DPAGE database (accession no. 0018). The addition of sequence information from ongoing fungal

proteome, genome and EST projects will increase the quality and efficiency of future proteomic studies on rusts and other fungal species. Knowledge of the ASR proteome will be useful for annotating the ASR genomic sequencing project that is currently underway as a joint project between the USDA-ARS, Department of Energy and Department of Homeland Security.

The authors thank Christine Stone for assistance in providing spore materials used in protein extractions. Commercial Endorsement Disclaimer – The use of trade, firm or corporation names on this page is for the information and convenience of the reader. Such use does not constitute an official endorsement of approval by the USDA Agricultural Research Service, NAL or BIC of any product or service to the exclusion of others that may be suitable. Equal Opportunity Statement – The US Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or a part of an individual's income is derived from any public assistance program. (Not all prohibited bases apply to all programs.) Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA's TARGET Center at (202) 720-2600 (voice and TDD). To file a complaint of discrimination, write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, SW, Washington, DC 20250-9410 or call (800) 795-3272 (voice) or (202) 720-6382 (TDD). USDA is an equal opportunity provider and employer.

The authors have declared no conflict of interest.

References

[1] Miles, M. R., Frederick, R. D., Hartman, G. L., Soybean Rust: Is the US Soybean Crop at Risk? *APSnet Features* 2003, online, DOI: 10.1094/APSnetFeature-2003-0603.

[2] Schneider, R. W., Hollier, C. A., Whitam, H. K., First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. *Plant Dis.* 2005, *89*, 774.

[3] Marchetti, M. A., Melching, J. S., Bromfield, K. R., The effects of temperature and dew period on germination and infection by urediniospores of *Phakopsora pachyrhizi*. *Phytopathology* 1976, *66*, 461–463.

[4] Bromfield, K. R., *Soybean Rust (Monograph*), Amer. Phytopathological Society, St. Paul, MN 1984.

[5] Bonde, M. R., Berner, D. K., Nester, S. E., Frederick, R. D., Effects of temperature on urediniospore germination, germ tube growth, and initiation of infection in soybean by *Phakopsora* isolates. *Phytopathology* 2007, *97*, 997–1003.

[6] Bonde, M. R., Nester, S. E., Berner, D. K., Frederick, R. D., Comparative susceptibilities of legume species to infection by *Phakopsora pachyrhizi*. *Plant Dis.* 2008, *92*, 30–36.

[7] Slaminko, T. L., Miles, M. R., Frederick, R. D., Bonde, M. R., Hartman, G. L., New legume hosts of *Phakopsora pachyrhizi* based on greenhouse evaluations. *Plant Dis.* 2008, *92*, 767–771.

[8] Edwards, H., Bonde, M. R., *Phakopsora pachyrhizi* host penetration strategy. *Phytopathology* 2008, *98*, S49.

[9] Ahern, K., Proteomic databases play increasing role. *Genet. Eng. News* 2004, *24*, 34–37.

[10] Ahn, I. P., Uhm, K. H., Kim, S., Lee, H. W., Signaling pathways involved in preinfection development of *Colletotrichum gloesporiodes*, *C. coccodes*, and *C. dematium* pathogenic on red pepper. *Physiol. Mol. Plant Pathol.* 2003, *63*, 281–289.

[11] Liu, Z. M., Kolattukudy, P. E., Identification of a gene product induced by hard-surface contact of *Colletotrichum gloeosporioides* Condida as a ubiquitin-conjugating enzyme by yeast complementation. *J. Bacteriol.* 1998, *180*, 3592–3597.

[12] Posada-Buitrago, M. L., Frederick, R. D., Expressed sequence tag analysis of the soybean rust pathogen *Phakopsora pachyrhizi*. *Fungal Genet. Biol.* 2005, *42*, 949–962.

[13] Markwell, M. A., Haas, S. M., Bieber, L. L., Tolbert, N. E., A modification of the Lowry procedure to simplify protein determination in membrane and lipoprotein samples. *Anal. Biochem.* 1978, *87*, 206–210.

[14] Perkins, D. N., Pappin, D. J. C., Creasy, D. M., Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999, *20*, 3551–3567.

[15] Pappin, D. J., Hojrup, P., Bleasby, A. J., Rapid identification of proteins by peptide-mass fingerprinting. *Curr. Biol.* 1993, *3*, 327–332.

[16] The Uniprot Consortium, The Universal Protein Resource (UniProt). *Nucleic Acid Res.* 2008, 36, D190–D195.

[17] Feder, M. E., Hofmann, G. E., Heat-shock Proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annu. Rev. Physiol.* 1999, *61*, 243–282.

[18] Nandi, D., Tahiliani, P., Kumar, A., Chandu, D., The ubiquitin-proteasome system. *J. Biosci.* 2006, *31*, 137–155.

[19] Vasara, T., Keranen, S., Penttila, M., Saloheimo, M., Characterization of two *14-3-3* genes from *Trichoderma reesei*: interactions with yeast secretory pathway components. *Biochimica Biophysica Acta* 2002, *1590*, 27–40.

[20] Warwar, V., Dickman, M. B., Effects of calcium and calmodulin on spore germination and appressorium development in *Colletotrichum trifolii*. *Appl. Environ. Microbiol.* 1996, *61*, 74–79.

[21] Cooper, B., Neelam, A., Campbell, K., Lee, J. *et al.*, Protein accumulation in the germinating *Uromyces appendiculatus* uredospore. *Mol. Plant Microbe Interact.* 2007, *20*, 857–866.

[22] Cooper, B., Garrett, W., Campbell, K. B., Shotgun identification of proteins from uredospores of the bean rust *Uromyces appendiculatus*. *Proteomics* 2006, *6*, 2477–2484.

[23] Rampitsch, C., Bykova, N., McCallum, B., Beimcik, E., Ens, W., Analysis of the wheat and *Puccinia triticina* (leaf rust) proteomes during a susceptible host–pathogen interaction. *Proteomics* 2006, *6*, 1897–1907.

